

Inter-rater reliability of *Health of Nations Outcome Scale* (HoNOS) among mental health nurses in Aceh

Sri Idaiani

Centre for Applied Health Technology and Clinical Epidemiology, National Institute of Health Research and Development, Indonesia Ministry of Health

Abstrak

Latar belakang: Alat ukur *Health of Nations Outcome Scale* (HoNOS) dapat digunakan untuk keperluan rutin dan menilai perkembangan status fungsional pasien gangguan jiwa di klinik maupun komunitas. Tujuan penelitian adalah melakukan uji kesepakatan (*agreement*) bagi perawat jiwa yang akan menggunakan alat ukur ini di Indonesia.

Metode: Lima puluh lima orang pasien psikosis yang sedang di rawat di Rumah Sakit Jiwa (RSJ) Provinsi Aceh dinilai secara bergiliran oleh 11 orang perawat jiwa menggunakan alat ukur HoNOS. Penilaian dilakukan pada bulan September 2011. HoNOS terdiri dari 11 pertanyaan. Kesesuaian di antara perawat terhadap masing-masing pertanyaan dinilai menggunakan koefisien *intra-class classification correlation* (ICC).

Hasil: Empat puluh dua persen pasien berusia 31-40 tahun, termuda berusia 24 tahun dan tertua berusia 59 tahun, mayoritas laki-laki, dan 38% mengalami psikosis selama 5-10 tahun. Tiga puluh enam persen perawat berusia >40 tahun sebagai penilai, separuhnya perempuan, dan 55% telah bekerja lebih dari 10 tahun. Nilai ICC untuk masing-masing pertanyaan secara umum baik (berkisar antara 0,8-0,9). Kesepakatan yang baik didapatkan di antara perawat Rumah Sakit Jiwa, perawat Puskesmas, maupun gabungan keduanya.

Kesimpulan: HoNOS memiliki *inter-rater agreement* yang baik dan dapat digunakan pada penelitian dengan setting yang sama. Untuk penggunaan di populasi yang lebih besar dan berasal dari daerah yang berbeda disarankan untuk melakukan uji reliabilitas serta validitas dengan jumlah sampel yang lebih besar. (*Health Science Indones 2011;2:53-7*)

Kata kunci: HoNOS, *agreement*, psikosis

Abstract

Background: The *Health of Nations Outcome Scale* (HoNOS) instrument could be used for routine purposes and assessing the functional status of the mental health patients in clinical and community settings. The objective of this study was to evaluate the agreement of the scale among mental health nurses who would use this tool in Indonesia.

Methods: Fifty five psychotic patients who were hospitalized at a mental hospital in Aceh were evaluated by 11 mental health nurses using the HoNOS instrument. The agreement between the nurses on each questionnaire item was evaluated using the *intra-class correlation* (ICC) coefficient.

Results: Forty-two percent of the patients were 31-40 years of age, the youngest was 24 and the oldest was 59, most of them were males, and 38% had psychosis for 5-10 years. Thirty-six percent of the nurses aged >40 years as raters, half of them were females, and 55% had worked for >10 years. ICC values were generally good (ranging from 0.8 to 0.9) among the mental health hospital nurses, as well as the community health centers nurses and a combination of both.

Conclusion: This instrument showed a good *inter-rater agreement* and could be used in future research with the same settings. For a wider use in different regions, it is recommended to test the reliability and the validity of the HoNOS in a larger study population. (*Health Science Indones 2011;2:53-7*)

Key words: HoNOS, agreement, psychosis.

Assessment of the effectiveness of a mental health intervention requires a measurement tool to assess the improvement of the outcomes of the mental disorders. Problems in the assessment of these outcomes include the absence of the following: a clear statement of the purpose of the intervention or treatment, clear outcome indicators, standardization of how to conduct such study, and the appropriate time intervals to assess the improvement. These findings from a study by Barbach and colleagues was cited by Trauer.¹ Assessment of the outcome of mental disorders using functional status is better than simply assessing the improvement of clinical symptoms.² Functional status assessment emphasizes more on improvement of one's ability in terms of functioning on a daily basis. If someone has a late disability, then the most important thing is the ability to function optimally.²

The Health of Nations Outcome Scale (HoNOS) is a measuring tool that can assess the development of functional status, especially for patients with mental disorders, both schizophrenia, psychotic and non-psychotic disorders, and others.^{1,3} This tool has been used in many countries, especially in the United Kingdom, Australia, and New Zealand. It has also been used in Canada, Germany, Italy, France and Norway. Assessments of the results have been published in numerous articles and publications.^{1,3-5} This tool was designed by Wing and colleagues in 1990.³ The HoNOS can be used for assessing the outcome of the mental disorders and other routine assessments. Previous studies show that the validity of the content, the test's reliability, and the inter-rater reliability were good.^{1,3-6} Other measuring tools that have been used in Indonesia are the Positive and Negative Symptoms Scale (PANSS) and the Brief Psychiatric Rating Scale (BPRS), which only measures the degree of symptoms alone; and the Global Assessment Functioning (GAF) to measure the patients' functional status.¹ However, GAF is widely used in clinical settings by psychiatrists to measure the improvement of the patients' functional status.² GAF was designed as a part of the fifth axis of the Diagnostic and Statistical Manual of Mental Disorders (DSM).

METHODS

This study was conducted in Aceh Province because Aceh is the only province in Indonesia that has community mental health programs that have been implemented in all districts and cities.

In the preparation phase, the HoNOS questionnaires were first translated from English into Bahasa Indonesia. Afterwards, they were back translated into English by different translators. This step was followed by a group discussion panel of researchers from NIHRD who produced a questionnaire measuring tool with content and translation in Bahasa Indonesia, in a way that it would at least be understandable by the community health center (CHC) nurses.

A pilot study was done to evaluate the HoNOS. Ten psychotic patients from six CHCs in the city of Banda Aceh and Aceh Besar district were assessed by six mental health nurses from the CHC. Each nurse assessed 1-2 patients using the HONOS based on the information contained in CHC medical records and interviewed patients and their families. Afterwards, the researchers had three sessions to discuss issues related to language, content, and the feasibility of this measuring tool. The Indonesian version of HoNOS, especially for the Aceh population, was finalized after the group discussion.

The inter-rater agreement of the Indonesian version of the HoNOS were evaluated among 11 nurses (six nurses who served at CHC in Banda Aceh and Aceh Besar district and five nurses who served at the Province of Aceh Mental Hospital). The mental health nurses involved in this study were nurses who run the mental health program at CHC and the selected CHCs have ample mental health patients. Nurses voluntarily joined the study. The flow of the research was described in Figure 1.

The nurses evaluated 55 psychotic patients who were being treated at the Province of Aceh Mental Hospital. The inclusion criteria includes residing in the city of Banda Aceh or Aceh Besar district and the patient was not expected to be discharged within 2 weeks from study inception. Patients in special wards (i.e. drug user or having physical comorbidity) were also excluded. Of 63 eligible patients, 55 patients were assessed by 11 mental health nurses between September 26 and October 4, 2011. Each rater assessed each patient.

Before assessing the patients, at the same time all nurses joined a 3-hour short course about how to use the measuring instrument in the hope they would have the same understanding of the meaning of the questions. The guideline was translated from the original, which was developed by Wing and colleagues.

The HoNOS consists of four dimensions of behavior, impairment, and social symptoms. The four dimensions

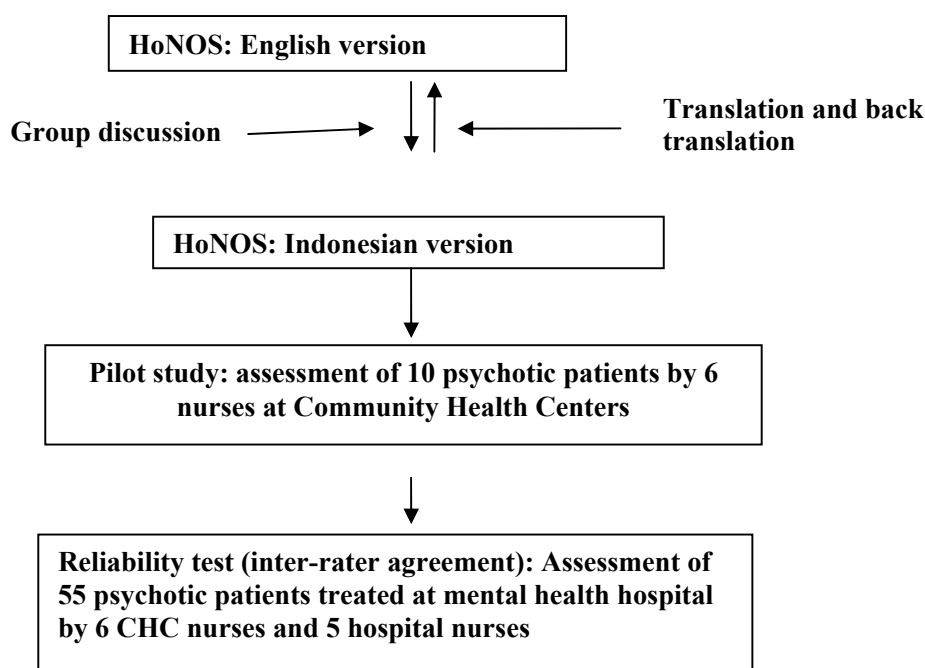


Figure 1. Flow of research

were translated into 12 items of questions. Items 1-3 were the domain of behavior; items 4-5 were included in the impairment domain; items 6-8 were in the symptoms domain; and items 9-12 were in the social domain. For each item, the score ranged from 0 to 4, with the following criteria: 0 = no problem; 1 = minor problem that does not interfere with the function; 2 = mild problem 3 = moderate to severe problem; and 4 = severe to very severe problem.. A value of 9 was assigned when the rater did not know or the condition of the patient did not fit the score criteria. Total scores and sub-total scores for each domain were calculated.

For this analysis, a patient's education was divided into three categories (low = no school or not completed primary school; medium = completed primary school or junior high school; high = senior high school or graduate schools). However, for the nurses only two categories (medium = high school; high = graduate schools) were used.

Inter rater-reliability analysis was performed by calculating intra-class correlation (ICC) coefficients using a two way mixed model and absolute agreement type. The average measure coefficients were reported. ICC is a good method to evaluate inter-raters agreement because ICC controls for measurement errors that might occur.⁷ Analysis was done using SPSS version 15.0.

Research permission was obtained from the Ministry of Home Affairs, while the ethics approval was obtained from the Ethics Committee of the National Institute of Health Research and Development (NIHRD), Ministry of Health of the Republic of Indonesia.

RESULTS

The characteristics of 55 patients are presented in Table 1. Patients were mostly male and more than 80% were between the ages of 24 and 39.

Thirty-six percent of the nurses were 40 or more years, half of them were female, and 55% had tenure of more than 10 years (Table 2). All nurses had attended the basic level of training on community mental health nursing (CMHN). Four nurses had an advanced course on CMHN.

Item number 2 of the HoNOS was not included in the analysis because there were only three patients who were reported to have a self harm problem. This low number might be explained because of good observation by hospital staff. Item number 3 (i.e. substance use) was also not analyzed because patients who used substances, including alcohol, were excluded from the study. Therefore, the analysis was just for 10 items.

Table 1. Some demographic characteristics of the patients

	n = 55	%
Age		
24-30 years	22	40.0
31-39 years	23	41.8
40-59 years	10	18.2
Gender		
Male	49	84.5
Female	6	10.3
Education		
Low	10	18.2
Medium	20	36.4
High	25	45.5
Duration of illness		
5-10 years	22	37.9
11-15 years	18	31.0
16-20 years	6	10.3
21-41 years	9	15.5

Nurses from the CHC were involved in the pilot assessment of the HoNOS, while the hospital nurses were not. The first group might know the measurement better than the second group. Therefore, the ICC were evaluated in three groups: the CHC nurses, the hospital nurses, and all nurses.

In general, the inter-rater agreement was good (ICC coefficients ranged from very good (0.81+), and good (0.61-0.80). Nurses from CHC were involved in the pilot study, while the hospital nurses were not. This might cause bias as the CHC nurses might have a better

understanding of the tool. However, the ICC coefficients between the nurse groups turned out to be comparable.

Table 2. Some demographic and employment characteristics of the nurses (raters)

	n	%
Age		
27-30 year	4	36.4
31-39 year	3	27.3
40-47 year	4	36.4
Gender		
Male	5	45.5
Female	6	54.5
Education		
Medium	2	18.2
High	9	71.8
CMHN course		
Never	3	27.3
Basic	3	27.3
Intermediate	1	9.1
Advance	4	36.4
Period of employment		
5-10 year	5	45.5
11-20 year	3	27.3
21-26 year	3	27.3
Work sites		
Mental hospital	5	45.5
Community health center	6	54.5

Table 3. Intra-class correlation coefficients of HoNOS items

No	HoNOS items*	Intra-class correlation coefficients*		
		CHC nurses (n = 6)	MH nurses (n = 5)	Total (n = 11)
1	Agression	0.93	0.97	0.96
4	Cognitive problem	0.92	0.97	0.94
5	Physical illness and disability	0.85	0.90	0.89
6	Hallucination and delusion	0.93	0.90	0.93
7	Depression	0.92	0.93	0.92
8	Other symptoms	0.92	0.94	0.95
9	Relationship	0.91	0.94	0.93
10	Activity Daily Living	0.92	0.95	0.91
11	Residential environment	0.91	0.88	0.84
12	Day time activity	0.84	0.61	0.82
	HONOS total score			
	1-10 items	0.91	0.92	0.94
	1-8 items	0.85	0.96	0.91

*The original items 2 and 3 were excluded because only 3 subjects had experienced self harm, while drug users were excluded from the study.

DISCUSSION

The reliability of the HoNOS among mental health nurses in this study was reliable. Previous studies in 293 patients in Manchester, 100 patients in Nottingham, and 50 patients in Geelong showed good reliability. Patients involved in those studies also had a diagnosis of mental disorder, which varied and was not limited to psychosis.¹⁻⁵

With these findings, the HoNOS could be used in future studies. Nevertheless, this research was done in Aceh because the province has a good community mental health program that was initiated after the tsunami disaster in 2004. In early 2005, the community mental health program was introduced and further developed in all health centers in Aceh province. At this moment mental health nurses could be found not only in mental hospitals, but also scattered in almost all CHCs in Aceh province. This is not the case in other provinces. Moreover, this study only assessed psychotic patients and the nurses who evaluated the subjects were mental health nurses. Therefore, if the HoNOS is going to be used in other places with different settings (e.g. other provinces and not limited to psychotic patients), the reliability of this measurement should be re-evaluated.

This study has several limitations. Sample size can affect the internal consistency, reliability and inter-rater reliability.⁸ Although this study was done in 55 psychosis patients, the findings were comparable to much bigger studies in Manchester and Nottingham^{1,3,5}. The validity, content, predictive, and diagnostic criteria of the translated HoNOS were not evaluated, although they were important aspects in trans-cultural epidemiology, because this study was a small part of a main study assessing the effectiveness of the community mental health program for psychotic patients at CHCs. Complete evaluation of these aspects should be considered in future studies.⁹

Acknowledgments

The author wishes to thank all subjects who willingly participated in this study. The author would also like to express her sincerest gratitude to Prof. Bastaman Basuki, Dr. Eva Suarhana, Dr. Muchtaruddin Mansyur, and Dr. Elisabeth Emerson for technical assistance in preparing this final draft.

REFERENCES

1. Trauer T, Callaly T, Hantz P, et al. Health of the Nation Outcome Scales Results of the Victorian field trial. *Br J Psychiatry*. 1999;174:380-8.
2. Cspike E, Wykes T. Global functioning scales. In: Thornicroft GT, M, editor. *Mental Health Outcome Measures*. 3 ed. London: RC Psych Pub; 2010.
3. Wing J, Beevor A, Curtis J, et al. Health of the Nation Outcome Scales (Ho NOS) research and development. *Br J Psychiatry*. 1998;172:11-8.
4. Pirkis J, Burgess P, Kirk P, Dodson S, Combs T, Williamson M. A review of the psychometric properties of the Health of the Nation Outcome Scales (HoNOS) family of measures. *Health Qual Life Outcomes* [serial on the Internet]. 2005;3.
5. Trauer T. The subscale structure of the Health of the Nation Outcome Scales (HoNOS). *J Ment Health* 1999;8:499-509.
6. Preston N. The Health of the Nation Outcome Scales: validating factorial structure and invariance across two health services. *Aust N Z J Psychiatry*. 2000;34:512-9.
7. Weir J. Quantifying test-retest reliability using the intra class correlation coefficient and the SEM. *J Strength Cond Res*. 2005;19(1):231-40.
8. Charter R. Study samples are too small to produce sufficiently precise reliability coefficients. *J Gen Psychol*. 2003;130:117-29.
9. Ommeren V. Validity issues in trans cultural epidemiology. *Br J Psychiatry*. 2003;182:376-8.